#5

A SHORT-TIME CORRELATOR FOR SPEECH WAVES

P. E. A. COWLEY

R. M. FANO

B. L. BASORE

TECHNICAL REPORT NO. 174

FEBRUARY 26, 1951

RESEARCH LABORATORY OF ELECTRONICS
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
CAMBRIDGE, MASSACHUSETTS

The research reported in this document was made possible through support extended the Massachusetts Institute of Technology, Research Laboratory of Electronics, jointly by the Army Signal Corps, the Navy Department (Office of Naval Research) and the Air Force (Air Materiel Command), under Signal Corps Contract No. DA36-039 sc-100, Project No. 8-102B-0; Department of the Army Project No. 3-99-10-022; and by the Robert Blair Fellowship of the London (England) County Council.

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
RESEARCH LABORATORY OF ELECTRONICS

Technical Report No. 174 February 26, 1951

A SHORT-TIME CORRELATOR FOR SPEECH WAVES

P. E. A. Cowley
R. M. Fano
B. L. Basore

Abstract

The equipment described is designed to analyze speech of a quality suitable for toll-telephone circuits. The analysis carried out is a thirteen-point approximation to the short-time, autocorrelation function which is presented as a display on a cathode-ray tube.

The design is discussed in some detail and circuit diagrams of the equipment are presented. Some experimental results are included of short-time correlation patterns of speech sounds, sinusoids, noise, and combinations of these waveforms.

The equipment is suitable for any application where autocorrelation or crosscorrelation functions of audio waveforms are desired to a moderate accuracy. Either short-time or long-time averaging may be accomplished.

A SHORT-TIME CORRELATOR FOR SPEECH WAVES

I. Introduction

Hitherto, speech sound waves have been analyzed mostly in the frequency domain. Thus, a sample of long duration may be characterized by a frequency power spectrum obtained from the time waveform by a generalized Fourier analysis. Such spectra have, in the past, proved useful in the design and evaluation of sound communication and reproduction systems. However, if the analysis is performed as a continuous process on samples of short duration, the resulting frequency pattern (short-time spectrum) changes with time and provides a dynamic representation of the characteristics of the waveform. Typical examples are visible speech (1) and vocoder (2).

Time-domain analyses of a waveform, analogous to the frequency analyses mentioned above, may also be performed. Such analyses are based on the autocorrelation function $\Phi_{11}(\tau)$, which is by definition the long-time average of the product of the speech waveform by the same waveform delayed τ seconds. Wiener (3) has shown that the autocorrelation function and the power spectrum of a time function are related to each other by a pair of Fourier transforms. A short-time autocorrelation function may also be defined by restricting the averaging operation to a short interval of time (4). As in the case of short-time spectra, the short-time autocorrelation function yields a pattern which changes with time. The information content of the speech wave found in this changing pattern is equivalent to that found in the corresponding frequency pattern and may well be readable when suitably presented in a manner similar to visible speech. There is the further possibility that some of the difficulty encountered in the reading of visible speech presentations might be overcome by this alternative method of analysis and presentation.

In the equipment described in this report, the averaging time has been made as long as possible to preserve the concept of autocorrelation function; at the same time, it has been kept sufficiently short to prevent the smearing of successive speech sounds.

The short-time autocorrelation function is displayed on an oscilloscope with the base line representing the delay τ and the vertical deflection representing the short-time correlation. A new pattern appears with each new phoneme as the speech progresses. This display is approximately the time-domain equivalent of an instantaneous time-varying display of the short-time power spectrum of the speech wave as it appears on the screen of a multichannel-spectrum analyzer (4).

II. The Design of the Correlator

A. Requirements. The correlator was designed to meet the following requirements:

1. perform continuous analysis of speech of telephone quality (approximately 300 cps to 4000 cps),
2. present the output on a cathode-ray tube in the form of a plot of correlation

as a function of time delay (τ),

3. respond adequately to a syllabic rate of at least 15 syllables each second.

B. Basic Design. The short-time crosscorrelation function may be defined by the analytic expression

$$\Phi_{12}(t, \tau) = \int_{-\infty}^{\infty} h(t-\sigma)f_1(\sigma)f_2(\sigma-\tau)d\sigma. \qquad (1)$$

The presence of the dependent variable t as an argument of the short-time crosscorrelation function serves to distinguish it from the more familiar long-time (3) crosscorrelation function $\Phi_{12}(\tau)$. $\Phi_{12}(t, \tau)$ becomes the short-time autocorrelation function $\phi_{11}(t, \tau)$ of $f_1(t)$ when $f_2(t) = f_1(t)$. The expression (1) is seen to be the convolution of a function h(t) with the product-time function $f_1(t)f_2(t-\tau)$. The function h(t) is called a scanning function, a window function, or an integration function. To perform ideal short-time averaging, the function h(t) should be rectangular, that is, zero for all negative t, unity for $0 \leqslant t \leqslant T$, and zero for all $t > T$. Any other function will give different weights to different parts of the product function. In a practical case, it is convenient to use narrow-band filtering, and in this instance h(t) is the impulse response of the filter. As is pointed out in reference 4, the particular case of a lowpass RC network $\left[h(t) = e^{-(t/RC)}\right.$ $\left.\text{for } t > 0\right]$ leads to a short-time autocorrelation function that is related to the frequency spectrum obtained by narrow-band filtering through a modified set of Fourier transforms. In practice, it is found that the shape of h(t) makes relatively little difference in the final result.

It is apparent from the expression $\Phi_{12}(t,\tau)$ that a device must be available to provide a delayed version of the waveform f_2. The delay device must be capable of delaying the function at least as long as the maximum delay for which $\phi_{12}(t, \tau)$ is desired.

The maximum value of τ for which $\Phi_{11}(t, \tau)$ must be evaluated is fixed by the lowest frequency component that is to be represented in the pattern. Since the autocorrelation function of a sinusoid is a cosine wave of the same period, a value of τ equal to at least half the period of the lowest frequency component is required. This assumes that a half-period is sufficient to permit the recognition of a particular frequency component. Reference 4, which covers this point in detail, leads essentially to the same result.

The number of discrete intervals into which the total maximum delay is divided sets the upper frequency limit as far as representation of the frequency components is concerned (4). As is recognized in sampling theory, it is necessary to take samples at least once each 1/2f sec to represent a waveform having frequency components of f cps or less. This corresponds to one sample per half-cycle of the highest frequency to be represented.

In this equipment a maximum delay of 1640 μsec, with eleven intermediate taps, will show a half-wave at 310 cps, and six full waves (each interval equal to a half-wave) at

3660 cps. The intervals are thus 137 μsec each. In this latter case, the sinusoid is represented by 13 alternate positive and negative points corresponding to the maxima and minima of the wave.

The required time delay may be obtained in a number of ways, among which one should mention lumped parameter networks, acoustic delay lines, magnetic tape or drum recording, and distributed (transmission line) parameter networks. The requirements are that the delaying device should have practically constant attenuation and delay over the range of speech frequencies to be delayed, and that it should further be readily tapped at the intermediate intervals of delay (a separate delay line for each value of delay might, of course, be used). A lumped constant bandpass delay line is used in the equipment described in this report.

III. An Outline of the Correlator

A block diagram is given in Fig. 1. The input signal or speech wave is used to modulate a 456-kc/sec signal in two distinct modulators. One is a conventional amplitude modulator by which a 456-kc/sec carrier is modulated 50 percent by the input speech wave. The other is a balanced modulator which produces a signal consisting only of the sidebands located symmetrically about the 456-kc/sec carrier. The AM signal is fed to the bandpass delay line from which outputs are taken at zero delay and at twelve successive taps, each corresponding to an increment of 137 μsec. The outputs are then separately demodulated by linear detectors, and each of the resulting delayed speech waveforms is applied to a multiplier strip as one factor of the product to be computed.

The balanced AM signal is fed commonly to all 13 channels as the second factor of the product to be computed. The multiplier itself consists of a receiver-converter tube, followed by a bandpass filter which is used to select the desired component of tube plate current. A phase-sensitive detector (P.S.D.) is used to demodulate the product component; the product function is integrated by a lowpass filter and then applied to the sequence gate or switching circuit. The half-power frequency of the lowpass filter is approximately 22 cps, which permits fluctuations at the syllabic rate to appear in the response but highly attenuates both 60-cps hum and all actual speech frequencies.

The switch connects in sequence each of the 13 channels to the vertical deflection plates of the oscilloscope as the trace takes the corresponding τ position. The rate of switching is 1000 times per second, a rate which produces a very nearly continuous trace on the screen, even for rapidly fluctuating, short-time correlation patterns.

Detailed descriptions of the blocks in Fig. 1 appear in later sections. However, a further word is pertinent in order to explain the choice of the particular type of multiplier used.

The product component of the plate current of the multiplier tube appears in a band of the spectrum in the immediate vicinity of the 456-kc/sec carrier. The advantages of having the product in this form are

a. The product is easily separated from the B+ circuit of the tube by means of transformers, even if the product itself is a constant. Otherwise (if a multiplier at audiofrequencies were used) a dc coupling circuit at high voltage level would be needed.

b. The product may be amplified in an ac amplifier until it becomes large compared to the drift associated with the following direct-coupled circuits.

c. The effect of second-harmonic distortion is not felt, because the frequency components caused by this distortion fall outside the limits of the filtering circuits.

It might well be said that the design of this multiplier trades the stability of a balanced germanium diode modulator for the unsatisfactory stability and distortion associated with direct multiplication and a direct-coupled amplifier.

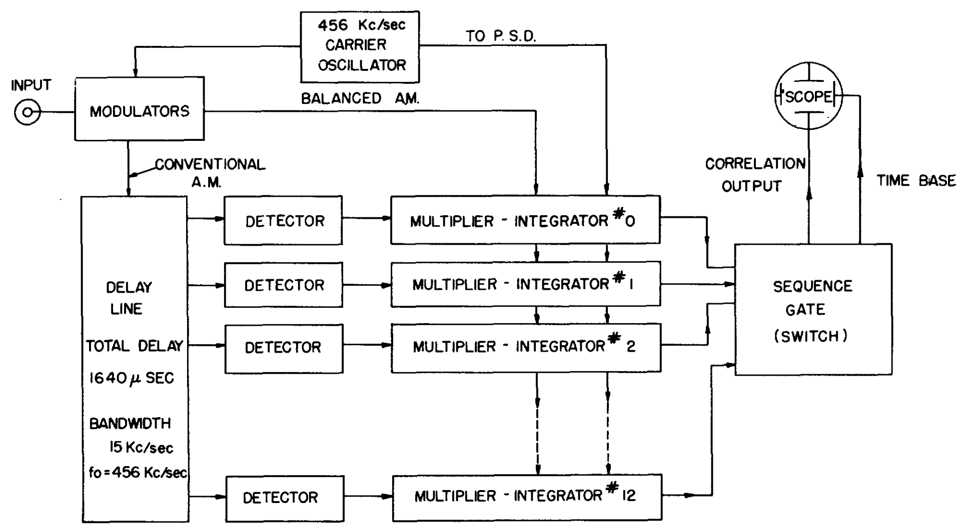


Fig. 1

Block diagram of correlator.

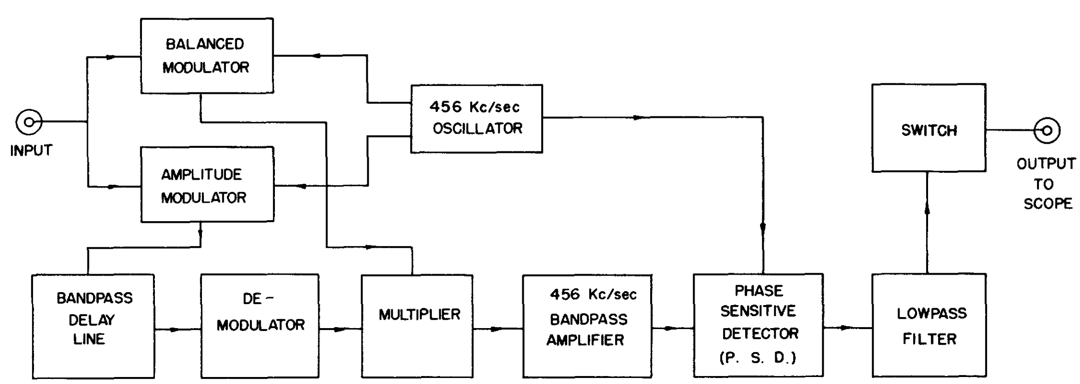


Fig. 2

Block diagram of one channel of the correlator using bandpass delay line.

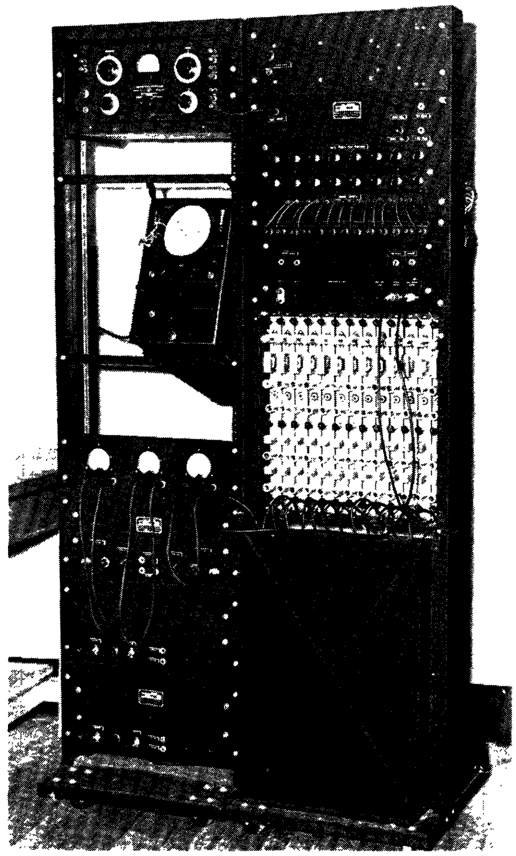


Fig. 3
Photograph of the correlator.

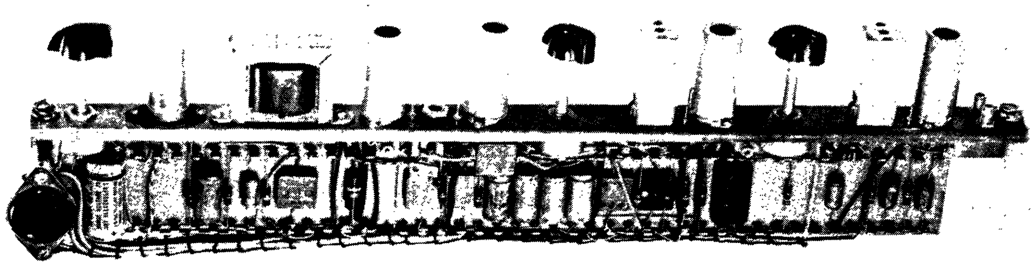


Fig. 4
Photograph of computing strip.

IV. Mechanical Layout

Figure 3 is a photograph of the equipment. The bandpass delay line is contained in the box at the bottom of the right-hand rack. Above it, arranged vertically, are the 13 computing strips corresponding to each of the 13 values of τ. Figure 4 shows the details of construction of one of the computing strips.

Above the computing strips is a modulator and distribution panel which also contains the carrier oscillator. Above this is the switching-circuit panel which is driven by the square-wave generator on the left-hand rack. Also, on the left-hand rack are the cathode-ray tube display (Dumont 208B modified for dc coupling) and the power supply panels.

V. General Circuit Description

The circuit is described sequentially insofar as that is possible. By reference to Fig. 2, the separate circuit descriptions can be combined into a description of the whole equipment.

The carrier frequency was chosen as 456 kc/sec because of the availability of i-f transformers. Miniature versions of these transformers and miniature tubes enable the channel equipment to be built on 1 1/4-inch chassis. Thirteen of these chassis are mounted side by side in a standard 19-inch rack.

VI. Speech Input and Modulator Circuit (Fig. 5)

Speech signals are applied from an unbalanced input circuit, via a speech-level control, and a 6-db, 10,000-ohm T attenuator to a ring modulator employing 1N34 germanium-crystal rectifiers. The carrier supply, 60 volts, 456 kc/sec, is reduced to approximately 6 volts at the modulator. Accurate modulator balance can be obtained with the three modulator balance potentiometers.

The 6BA6 amplifier tube is coupled to the modulator by a bandpass filter consisting of an i-f transformer with a bandwidth of 14 kc/sec. This filter is necessary for the elimination of unwanted modulation products, especially those involving the third harmonic of the carrier oscillator frequency. The delay introduced by this filter (approximately 11 μsec) is compensated by an approximately equal delay in the amplitude modulator. The latter single-tube modulation, however, does not require filtering of the output, since that function is performed by the bandpass delay line. Some additional delay compensation is accomplished by the RC network which applies the audio voltage to the af grid of the modulator tube. The 456-kc/sec output is modulated approximately 50 percent by the input speech signal of 1-volt (rms) amplitude.

VII. Computing Strip Circuit (Fig. 6)

Each computing strip contains the entire circuit for one channel of the correlator, that is, detector, multiplier tube, product filter and amplifier, phase-sensitive detector, and output filter.

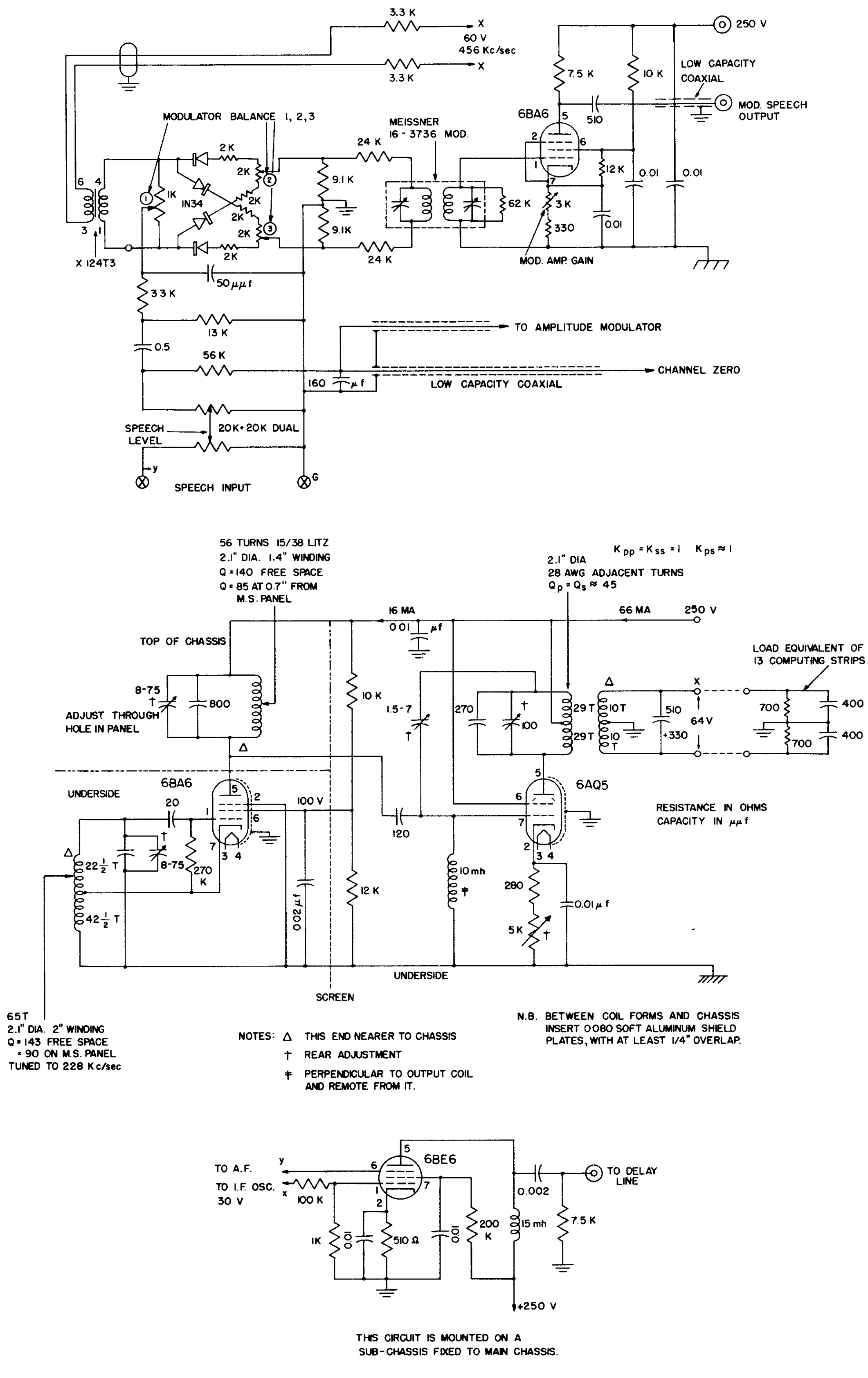
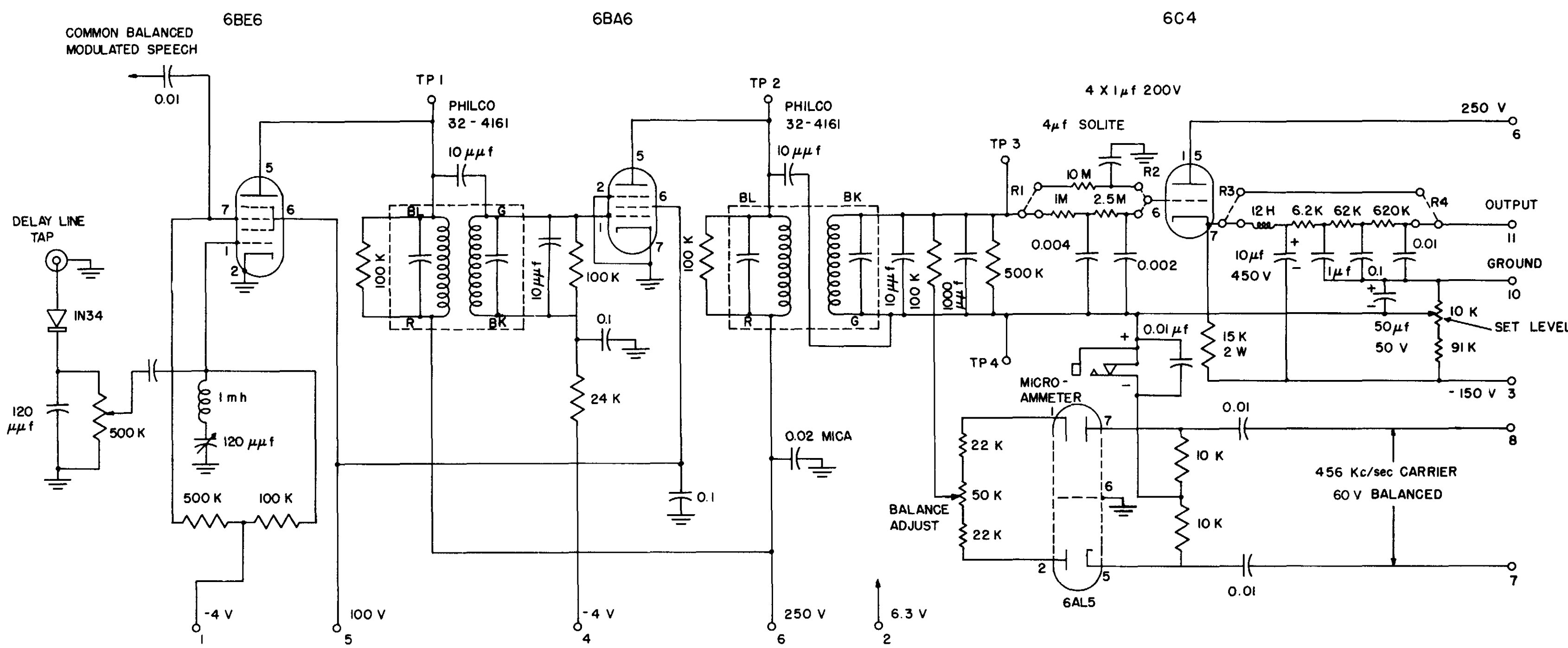


Fig. 5

Modulator distribution panel (distribution circuit omitted).



◎ MINIATURE COAXIAL CONNECTOR

o CONNECTION TO DISTRIBUTION PANEL VIA CABLE AND 11-WAY CONNECTOR

R1---R4 STRAPPING POINT

TP TEST POINT

Fig. 6

Speech correlator computing strip circuit.

The multiplier tube (6BE6) is of the converter type; it has a plate current that is proportional to the product of the two signal grid voltages. If $A + f(t) \sin \omega t$ is the bias plus the speech-modulated (with suppressed carrier) signal which is applied to one signal grid, and $B + f(t-\tau)$ is the bias plus delayed speech which is applied to the other signal grid, the plate current has a component of the form

$$I_p \approx AB + A\, f(t-\tau) + B\, f(t) \sin \omega t + f(t)\, f(t-\tau) \sin \omega t.$$

Of these product terms, only the third and fourth can pass through the product filter and amplifier (6BA6). The third term is eliminated after detection in the output lowpass filter. The level of this third term is some 18 db to 20 db above that of the fourth term which represents the desired product.

The product filter, which consists of the 456-kc/sec i-f transformers in the product amplifier, is adjusted to have approximately $\pm$20-kc/sec flat response, symmetrical about 456 kc/sec.

Demodulation of the product is carried out in a half-wave, phase-sensitive detector. The duo-diode 6AL5 conducts on alternate half-cycles of the carrier voltage to "switch" the secondary of the i-f transformer across the load of 500 K and 1000 $\mu\mu$F. The level of the desired product which occurs at the load resistor is approximately 2 volts, but the undesired term (delayed speech) is at a level of approximately 20 volts. It is because of the high level of this undesired term (which is eliminated later in the circuit) that the high carrier voltage (60 volts) is used. A low carrier voltage would result in distortion in the demodulator. The balance adjustment enables the elimination of any dc output due to unbalance of the demodulator.

The output filter must provide sufficient attenuation in the speech-frequency range to make the level of the speech negligible compared with the level of the averaged product. The lowest frequency components of speech are in excess of 100 cps, and their level must be reduced 40 db below the level of the desired output. On the other hand, the speech is already +20 db above the desired output, so that an attenuation totaling 60 db is required at 100 cps. The correlator output must vary at syllabic rates up to at least 15 cps, so that the output filter should cut off at a frequency of not less than 15 cps. The required characteristic is readily obtained by means of the two-section RC filter preceding the 6C4 cathode follower and the four-section RLC filter following it. The inductance in this filter serves the purpose of maintaining a flat response up to 15 cps. Although reference 4 indicates that an RC filter might be desirable in order to obtain a final output with a known relation to the short-time frequency spectrum, the attenuation requirements have necessitated the use of the filter described.

By reconnection at the strapping points R1, R2, R3 and R4, the output filter is changed to an RC lowpass circuit with a time constant of 40 sec. With this connection, the equipment yields a closer approximation to the long-time correlation function of the input waveform.

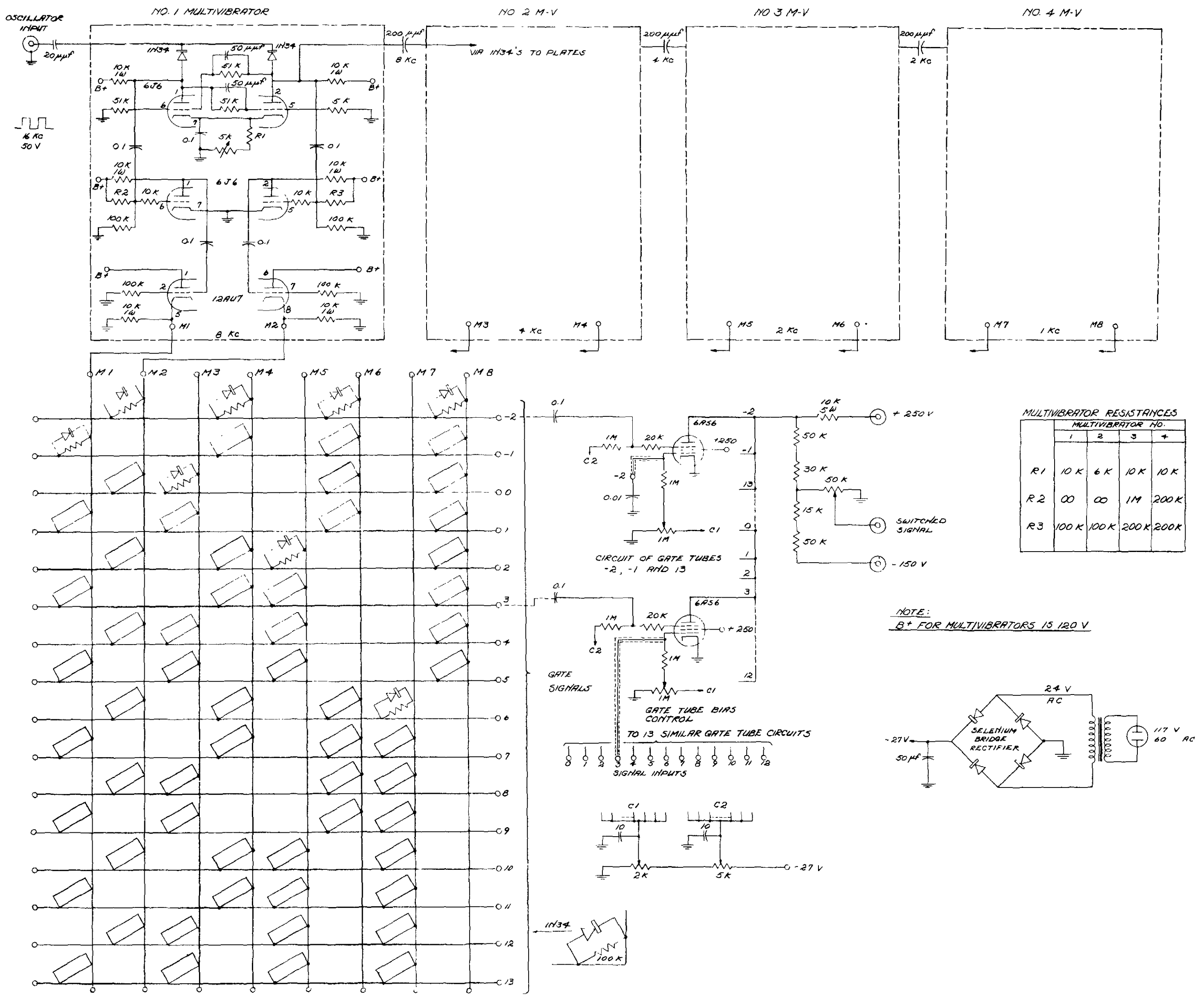


Fig. 7
Switching panel for short-time correlator.

VIII. The Switching Circuit (Fig. 7)

Associated with each channel is a gate tube which connects that channel in its turn to the deflection plates of the cathode-ray tube. The switching circuit works on a binary system, and has 16 gate tubes which could accommodate 16 channels. Since there are only 13 channels, the three remaining gate tubes are available for use as reference markers. The zero correlation position on the cathode-ray-tube display is indicated by two markers preceding the correlation pattern and one following it. A marker gate tube is shown in the circuit as the upper 6AS6 tube. The lower 6AS6 tube is the gate tube associated with channel 3. The gate tubes associated with the other channels are similar. The plates of all 16 gate tubes are commoned and dc-coupled to the cathode-ray-tube display via the switched-signal output socket.

All gate tubes are normally cut off by suppressor grid bias except for the one tube which, at any one time, has a positive gate signal. This tube causes the common plate circuit to take up a potential determined by the control grid potential. The control grid potential is derived from the output filter of the appropriate channel. Gate signals of +20-volt amplitude and 1/16, 000 sec in duration, with a repetition rate of 1000/sec, are generated in the crystal resistance matrix. The method of obtaining a switching system with a crystal resistance matrix has been described in detail elsewhere (5). The matrix is driven by balanced square-wave sources of 8 kc/sec, 4 kc/sec, 2 kc/sec, and 1 kc/sec. The sources consist of a chain of flip-flop stages (with associated amplifiers and cathode followers) driven by a 16-kc/sec, 50-volt square-wave generator.

Consider the gate signal (No. 3) for the lower 6AS6 tube shown in the circuit. The horizontal bar of the matrix labelled "gate signal No. 3" is connected by crystals to the four vertical bars, M1, M4, M5, and M8. A positive gate signal is possible only when all four of these points are positive. Such a condition occurs for 1/16, 000 sec during each repetition period of 1/1000 sec, and similarly for the other gate signals. Gate signal No. 13 is used to synchronize the linear sweep of the cathode-ray-tube trace.

IX. Carrier Oscillator (Fig. 8)

This is a two-tube oscillator which provides an accurately balanced-to-ground 60-volt output for the modulator and phase-sensitive detectors. Air-core coils are used. In order to maintain a reasonable Q value when the coils are mounted on a mild steel chassis, it was found necessary to insert aluminum shielding washers between the coil forms and the mild steel chassis.

The carrier oscillator is mounted on the modulator and distribution panel. The distribution is shown in Fig. 9.

X. The Bandpass Delay Line

A bandpass, lumped-constant delay line was chosen over the other possible types of delay mechanism for the following reasons:

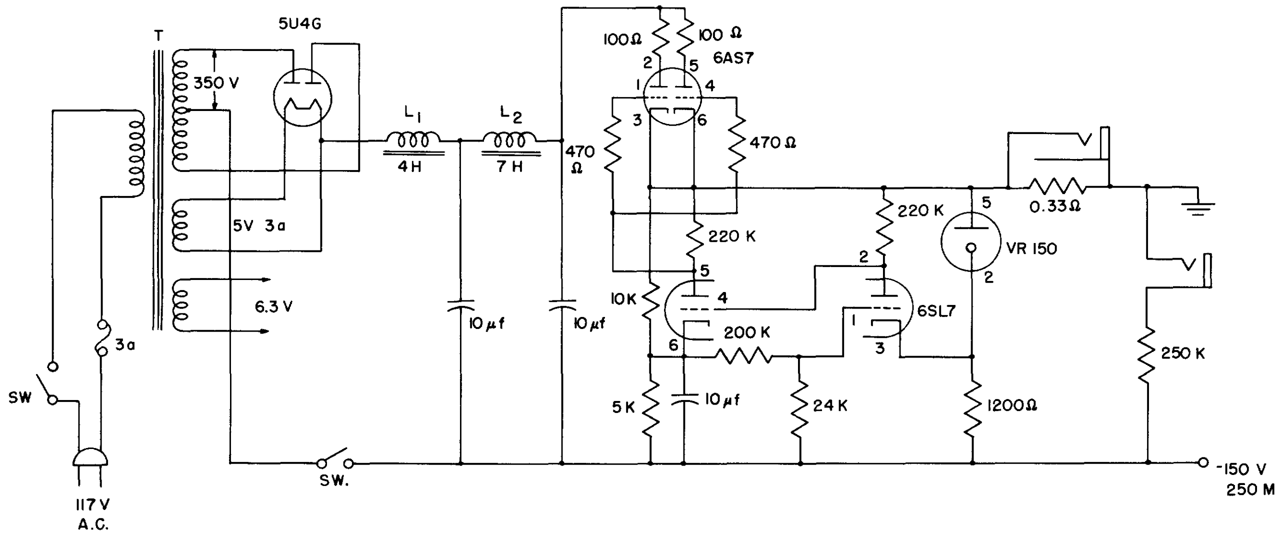


Fig. 8
The -150 volt power supply.

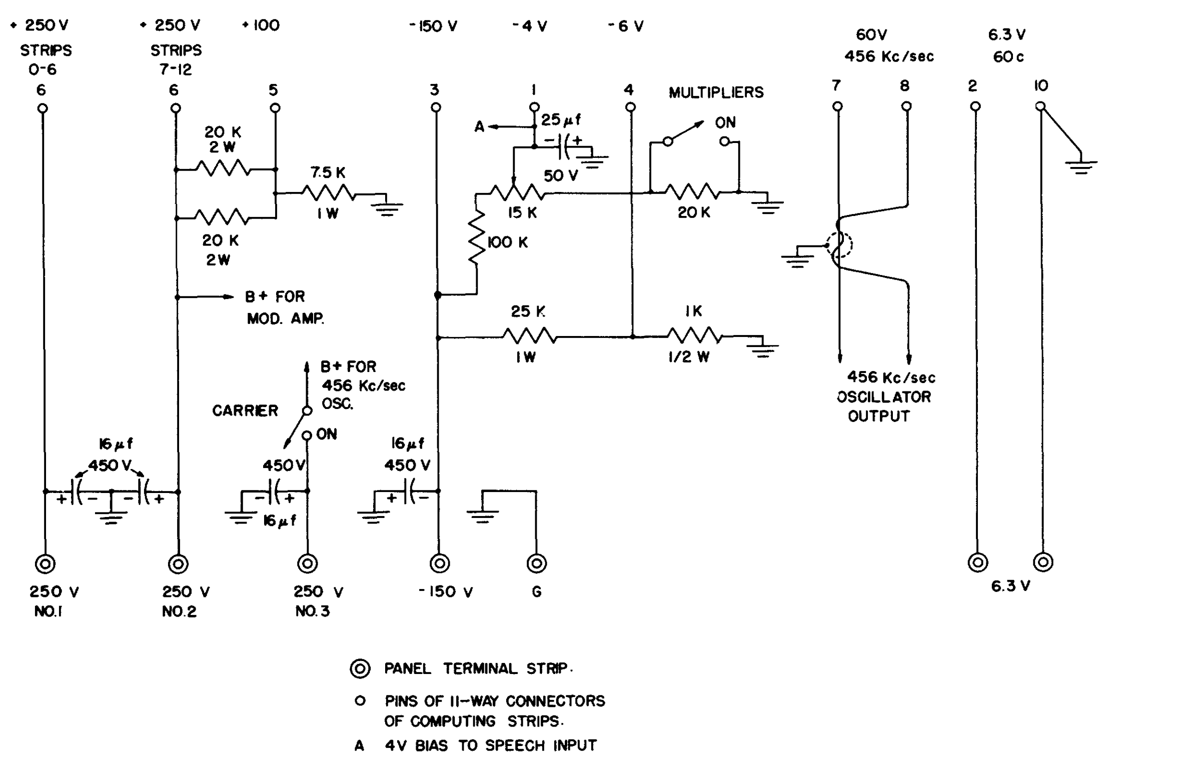


Fig. 9
Modulator and distribution panel showing distribution circuit.

a. Supersonic and acoustic delay lines are considerably more expensive when a relatively large number of taps are required.

b. Magnetic drums are not suitable for the short delay intervals required.

c. Inexpensive pi-wound, air-core coils may be used; a lowpass type of line would require expensive powdered-iron toroidal coils

d. Losses in coils of a bandpass delay line are substantially constant over the frequency band. In the lowpass delay line, the losses vary greatly over the frequency band, so that it is difficult to obtain even approximately uniform attenuation.

e. The bandpass delay line may be aligned by means of small, continuously variable, trimmer capacitors; a lowpass type of line could not be aligned in such a convenient manner.

The disadvantage of the bandpass line lies in its greater attenuation. The disadvantage of additional equipment, that is, a modulator and demodulator for the signal which is to be delayed, is seemingly minor when compared to the advantages listed above.

Line attenuation is easily overcome by amplifiers at intervals along the line. The use of amplifiers might at first appear to be a disadvantage in that the line then becomes unidirectional, but this in no way impairs its function in the correlator. Actually, the presence of amplifiers at intervals along the line is a great advantage in that when the line is being tuned, a part of the line between amplifiers does not interact with other parts. Although component coils, capacitors, and couplings are accurately adjusted before assembly of the line, tuning is still necessary to obtain uniform attenuation. Owing to interaction between sections of the line, the difficulty of tuning increases rapidly with the number of sections in each part of the line. From experimental work, nine sections in each part were found to represent a practical limit.

The line has a total of 324 sections in 36 parts of 9 sections each. Following each part is an amplifier (Fig. 10) with a gain of approximately 6 db and an output impedance of 7500 ohms, to match the impedance of the line part which follows. Gain adjustment enables the gain to be adjusted to equal the loss of the preceding line part. The line is tapped each three parts to obtain an increment of 137-μsec delay per tap.

Each line part (Fig. 11) comprises nine sections in one subassembly (7 center sections and 2 end sections). In order to be able to use the same coils and capacitors in the end sections as in the center sections, the designed line impedance of 15,000 ohms is reduced to 7500 ohms in the end sections by an impedance transformation.

The nine-section filter consists of nine tuned circuits, inductively coupled. The coils are pi-wound, arranged axially on a bakelite tube, and are accurately spaced by means of bakelite spacing sleeves between the coils. In order to reduce coupling between nonadjacent coils to negligible proportions, the coil assembly is mounted axially in a short section of brass tubing. The tubing acts as a waveguide below its critical frequency and causes the magnetic field to decay exponentially with distance.

The bandwidth of the nine-section filter is approximately 40 kc/sec. However, with 36 such filters in series to form the complete delay line, the bandwidth, even with careful tuning, is reduced to 14 kc/sec or 16 kc/sec. The requirement of the present correlator is for a bandwidth of 8 kc/sec (to accommodate a double sideband of speech having a maximum frequency of 4 kc/sec). The attenuation and phase characteristics of the complete line are given in Fig. 12. The attenuation is substantially uniform (within 0.2 db), and the delay is substantially constant (within 2 percent) in the frequency band of interest.

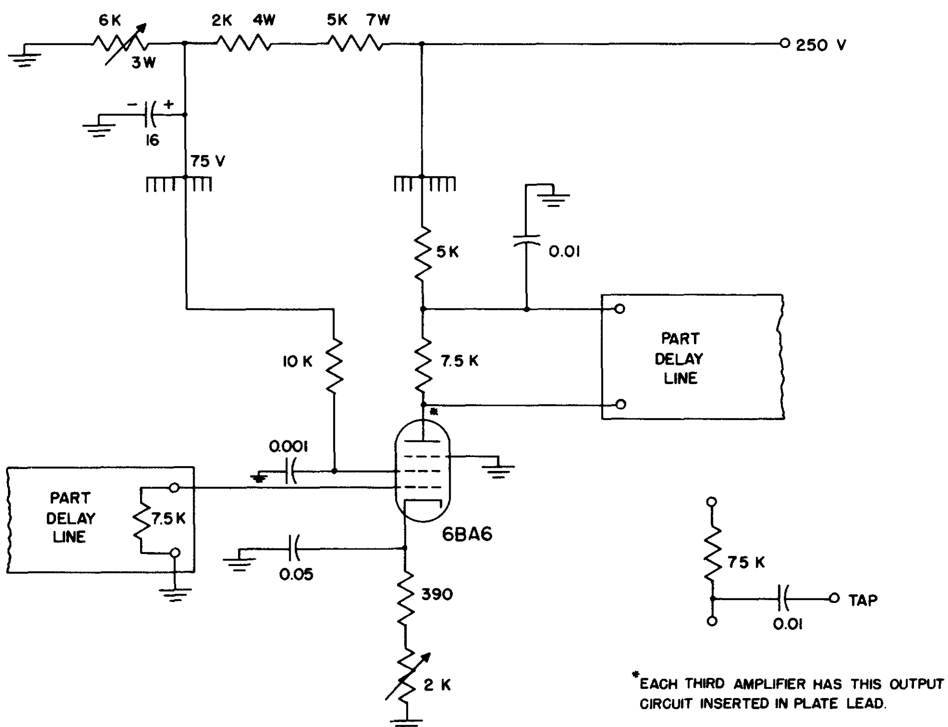


Fig. 10
Delay-line amplifiers.

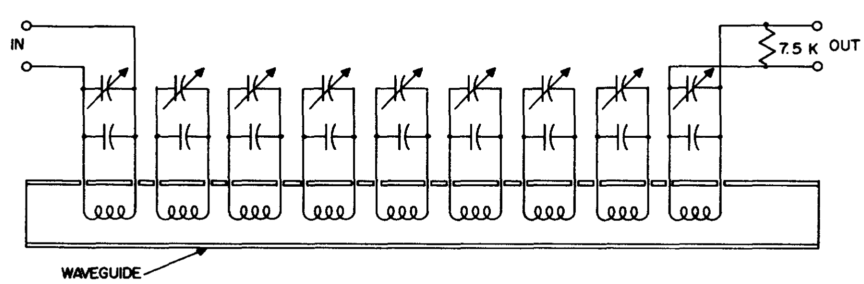


Fig. 11
One 9-section part of bandpass delay line.

XI. Calibration and Operation of the Equipment

At least two hours warm-up time must be allowed before calibration of the equipment can be accomplished, and certain preliminary steps must then be taken. First, it is necessary to adjust the balanced modulator accurately to remove any traces of carrier in its output. This is most readily done by first adjusting balance control 1 (see Fig. 5) until no carrier appears across the input terminals. Then controls 2 and 3 are adjusted until no 456-kc/sec signal appears at the output when the input is short-circuited. The three controls interact somewhat, and sequential adjustment is necessary to complete the balance.

Next, with the carrier off, the "set level" control on the top of the computing strip (see Fig. 6) is adjusted in each channel to produce a flat trace on the oscilloscope. The position of the trace is determined by the "-2," "-1," and "13" positions of the trace, which are adjusted to the middle of the linear response portion of the 6AS6 gate tubes.

Then, with the carrier turned on, the phase-sensitive detectors are balanced by the appropriate potentiometers. Balance is determined when, again, a flat trace obtains.

Finally, with a one-volt (rms) variable frequency voltage applied, starting at the top frequency (for which channel 1 shows maximum negative deflection – approximately 3680 cps) and decreasing the frequency, each channel gain control is adjusted so that equal maximum negative deflection occurs as each channel passes through a minimum.

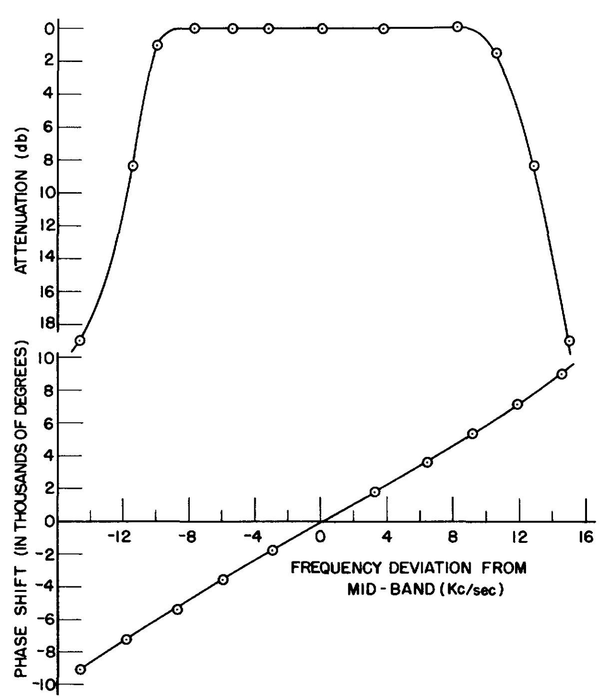
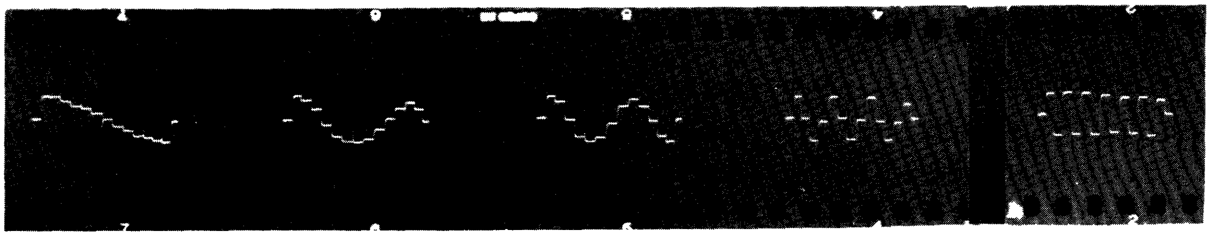


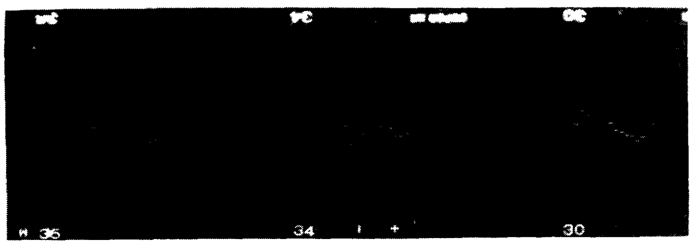
Fig. 12
Attenuation and phase characteristic of the delay line.



(a) 325 cps (b) 650 cps (c) 960 cps (d) 1880 cps (e) 3760 cps

Fig. 13

Correlation patterns of sinusoids of
325 cps, 650 cps, 960 cps, 1880 cps, and 3760 cps, respectively.



(a) (b) (c)

Fig. 14

Correlation pattern of two sinusoids applied simultaneously and separately.

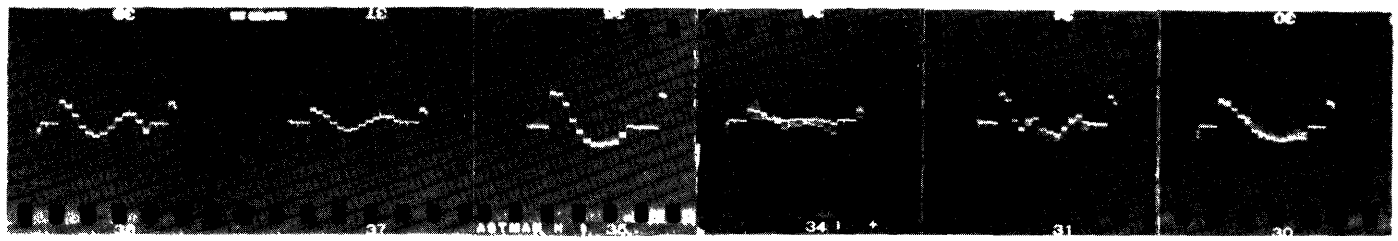
The magnitude of this deflection should be approximately one inch. Channel zero is adjusted to fall into a smooth curve when channel 12 is showing maximum negative deflection. The equipment is now ready for operation.

XII. Correlation Patterns

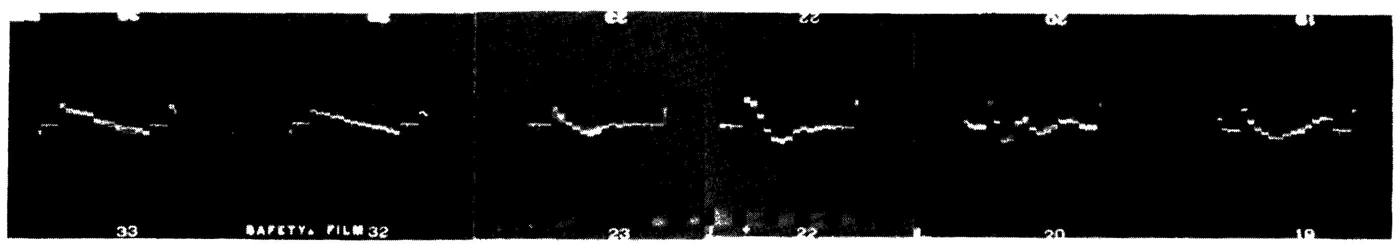
A. Sinusoids. A check on the operation of the correlator and an indication of the representation is obtained by examining the correlation patterns obtained for sinusoidal inputs. Patterns for frequencies of 325 cps, 650 cps, 960 cps, 1880 cps, and 3760 cps* are given in Fig. 13. According to theory (6), the correlation pattern of a sinusoid is a cosine wave of the same period, and the figures show that the correlator gives the best 13-point approximation to the cosine curve. Regardless of their relative phase, two sinusoids applied simultaneously result in a correlation pattern which is the sum of two cosine waves having the same periods as the input sinusoids (6). This effect is shown in Fig. 14.

---

*The frequencies listed are those read on the audio oscillator dial.



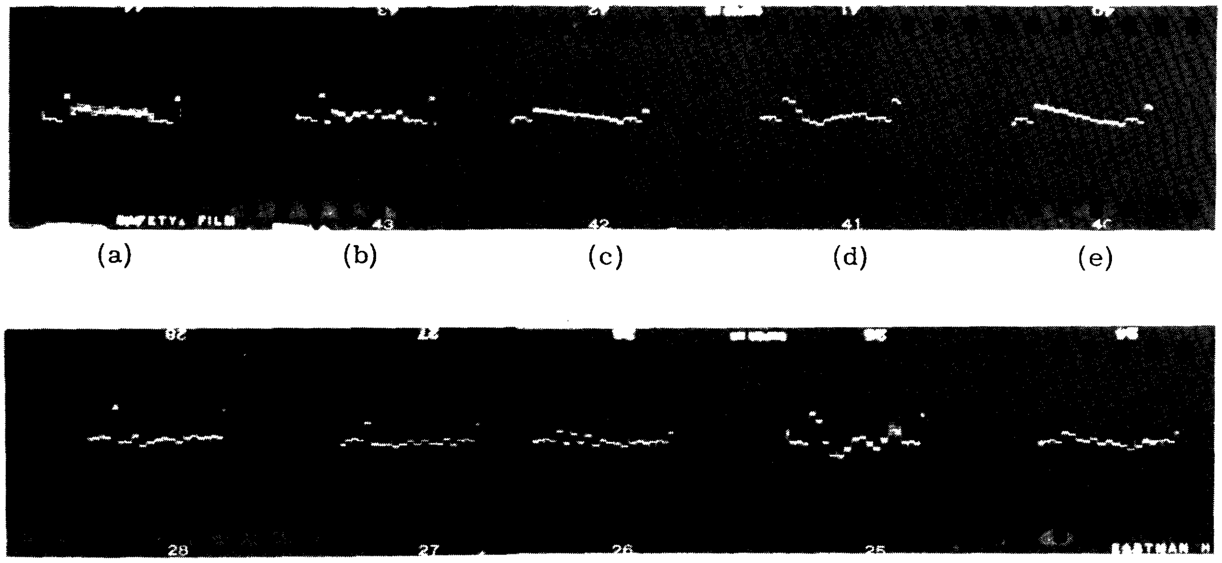
(a) ask (b) all (c) for (d) vote (e) church (f) judge



(g) she (h) azure (i) for (j) vote (k) church (ℓ) judge

Fig. 15

Short-time correlation patterns of speech sounds: (a) - (h) ordinary speech, (i) - (ℓ) differentiated speech.



(a) (b) (c) (d) (e)

(f) (g) (h) (i) (j)

Fig. 16

Short-time correlation patterns of speech sounds:
(a) - (e) ordinary speech, (f) - (j) differentiated speech.

XIII. Speech

A. Short-Time Correlations. Correlation patterns similar to those discussed above, not only can be, but have been, obtained by the use of long-time correlation methods (7, 8). The correlation pattern of short speech sounds, however, can only be obtained with a short-time correlator. Correlation patterns of sustained speech sounds have been made by other investigators (9) and agree fairly well, on the whole, with patterns obtained from the short-time correlator.

Short-time correlation patterns of speech sounds are given in Figs. 15 and 16 for both ordinary speech and differentiated speech. In the sonograph equipment for obtaining visible speech patterns, differentiation is used to assist in revealing the higher frequency vowel bars. It was anticipated that in the correlation pattern the differentiation would reveal a more detailed pattern. Unfortunately, there is some doubt as to the validity of the correlation patterns of differentiated speech, because with differentiated speech some considerable speech power appears at frequencies above the 4000-cps upper design limit of the correlator. These frequencies, when applied to the correlator, appear as frequencies below the upper frequency limit (at which each interval of delay corresponds to a half-wave) by precisely the amount that they are actually above that frequency. For example, twice the upper frequency limit is such that each delay interval corresponds to a complete cycle, and a pattern typical of direct current is obtained.

B. Long-Time Correlations. Using an averaging period of 40 sec in the computing strip, it was found that the correlation pattern was stationary for speech not containing

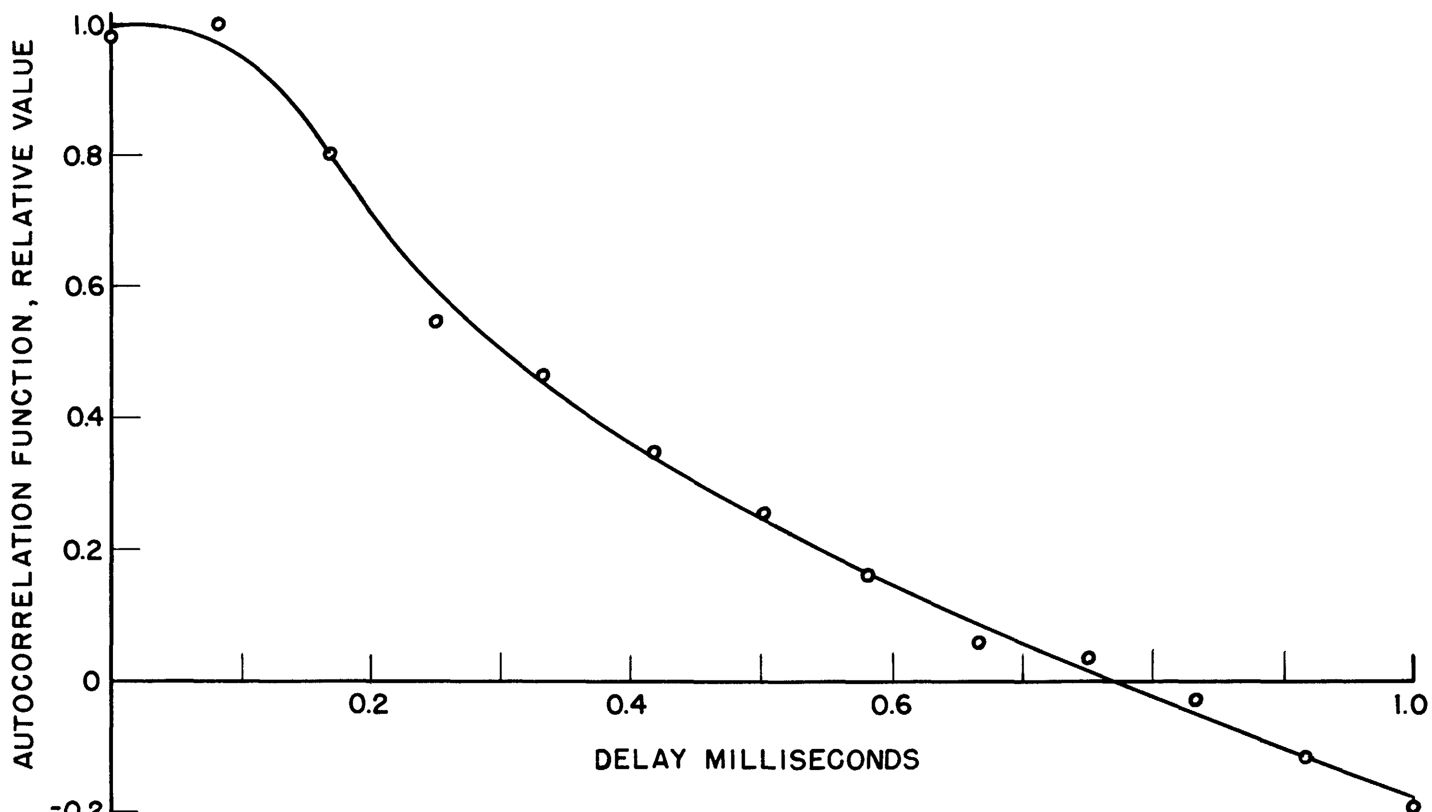


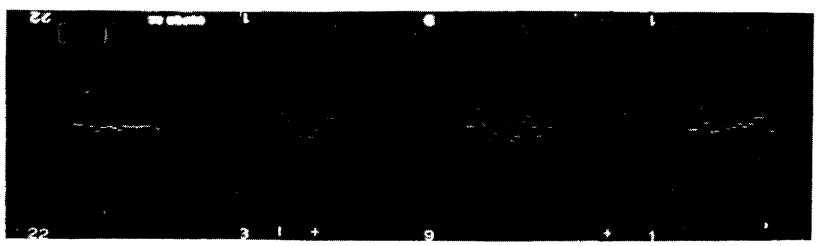
Fig. 17
Long-time correlation pattern of speech.

pauses of undue length. Figure 17 shows the curve obtained by a point-by-point method, using only one computing strip and a delay line of somewhat shorter length than the line described above. The curve has not been checked since completion of the correlator, but it compares favorably with curves obtained with other equipment (7).

XIV. Noise

A. Short-Time Correlation of Noise. The correlation pattern produced by noise that has an approximately uniform distribution in the frequency range 0-3000 cps is shown in Fig. 18a. It may be described as a δ-function at $\tau = 0$.

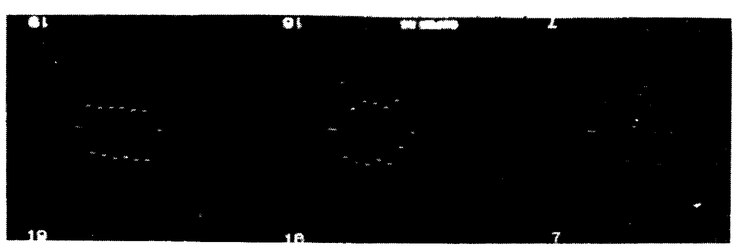
The correlation pattern of noise modified by a tuned circuit is a damped cosine wave having a period equal to that of the tuned circuit, and a damping factor that is inversely proportional to the Q of the tuned circuit (10, 11). In Fig. 18 (b, c, and d) is shown the correlation patterns of noise through tuned circuits of high, medium and low Q.



(a) (b) (c) (d)

Fig. 18

Correlation pattern of noise having: (a) wideband distribution, (b) distribution modified by tuned circuit of high Q, (c) distribution modified by tuned circuit of medium Q, (d) distribution modified by tuned circuit of low Q.



(a) (b) (c)

Fig. 19

Correlation patterns of sinusoidal signals and noise: (a) 3800 cps sinusoid and no noise, (b) 3800 cps sinusoid and wideband noise, (c) 960 cps sinusoid and noise modified by a tuned circuit of medium Q factor.

B. Short-Time Correlation of Noise and Sinusoidal Signals. In Fig. 19 (a and b) is shown the correlation pattern of a 3800-cps sinusoidal signal without noise and with wideband noise (noise and signal of approximately equal rms amplitude). The pattern is not quite stationary, but it is easy to recognize the presence of the sinusoidal signal.

Figure 19c shows the correlation pattern of a 960-cps sinusoidal signal together with medium-band noise at a somewhat lower level through a single tuned circuit. Recognition of the signal is not so easy because of the contribution of the noise to the whole correlation pattern. However, the three half-wavelengths of the 960-cps signal can be identified, although not nearly so readily as in the case of a 960-cps noise-free signal (compare Fig. 13).

XV. Speech and Noise

Experimental observation of correlation patterns of noise and speech sounds were not made, and there is no indication that speech sounds will be readily recognizable in the presence of noise. However, we may reasonably expect that wideband noise (which is characterized in the correlation pattern by a δ-function at $\tau = 0$) will not greatly hinder recognition of the vowel sounds, but that such noise would add considerably to the difficult problem of recognition of certain consonants. It has not been possible at the time of submitting this report to conduct the extensive experiments which would be required to compare the presentation of the short-time autocorrelation function with the presentation of visible speech.

XVI. Future Application of the Short-Time Correlator

Apart from further work with speech, there is the possible application of the short-time correlator to the detection of nonuniformly repetitive signals in noise.

In fact, the equipment seems useful in any case where long-time or short-time correlation functions are desired to an accuracy of approximately ± 5 percent, and where the random time series occupies a spectrum approximately that of speech or can be scaled to fit into this requirement.

References

1. Ralph K. Potter et al: Visible Speech, J. Acoust. Soc. Am. 18, 1-89, 1946
2. H. Dudley: Remaking Speech, J. Acoust. Soc. Am. 11, 169, 1939
3. N. Wiener: Acta Mathematica 55, 273, 1930
4. R. M. Fano: Short-Time Autocorrelation Functions and Power Spectra, J. Acoust. Soc. Am. 22, 546, 1950
5. L. L. Kilpatrick: A High Capacity Matrix-Commutated Telemetering System, Master's Thesis, Dept. of Electrical Engineering, M.I.T. 1948
6. Y. W. Lee: Class Notes, Subject 6.564, Dept. of Electrical Engineering, M.I.T.
7. T. P. Cheatham: An Electronic Correlator, Technical Report No. 122, Research Laboratory of Electronics, M.I.T. March 1951
8. H. E. Singleton: A Digital Electronic Correlator, Technical Report No. 152, Research Laboratory of Electronics, M.I.T. Feb. 1950
9. Visual Message Presentation, Northeastern University Quarterly Progress Report No. 5, Contract No. AF19(122)-7, Aug. 8, 1950
10. Y. W. Lee, T. P. Cheatham, J. B. Wiesner: Application of Correlation Analysis to the Detection of Periodic Signals in Noise, Proc. I.R.E. 38, 1165, 1950
11. Y. W. Lee, C. A. Stutt: Statistical Prediction of Noise, Technical Report No. 129, Research Laboratory of Electronics, M.I.T. July 1949